8-2020

Analyzing the Fractal Dimension of Various Musical Pieces

Nathan Clark

ANALYZING THE FRACTAL DIMENSION OF VARIOUS MUSICAL PIECES

NATHAN CLARK

A thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Industrial Engineering

Thesis Advisor: Justin R. Chimka, Ph.D.
Thesis Reader: Chase E. Rainwater, Ph.D.

Abstract:

One of the most common tools for evaluating data is regression. This technique, widely used by industrial engineers, explores linear relationships between predictors and the response. Each observation of the response is a fixed linear combination of the predictors with an added error element. The method is built on the assumption that this error is normally distributed across all observations and has a mean of zero. In some cases, it has been found that the inherent variation is not the result of a random variable, but is instead the result of self-symmetric properties of the observations. For data with these characteristics, fractal analysis can be used to explain the variation. There has been evidence from previous work that musical pieces have to some degree a fractal structure, but there remains to be more work done on performing fractal analysis to musical pieces. In this research, a computationally efficient method of performing fractal analysis on time-series data is applied to a musical recording. It is then determined whether this fractal dimension is a suitable measure to distinguish between musical genres.

Awknowledgements:

My top thanks goes to my dad, who has constantly inspired me from the very start, far more than I ever realized. Thanks also to this department. It is something special here on campus. I have not only been taught, but been loved by the many faculty and staff during my time here. Thanks specifically to my advisor, Dr. Chimka, who has made this work and my success possible.

Date: July 30, 2020.

Contents

1. Introduction to Fractals 3
    1.1. The Cantor Set 4
    1.2. The Koch Curve 5
    1.3. The Sierpinski Triangle 6
2. The Fractal Dimension 8
    2.1. The Hausdorff Dimension 8
    2.2. Box-Counting and Other Fractal Dimensions 10
3. Measuring the Fractal Dimension of Time-Series Data 12
4. Measuring the Fractal Dimension of Music 14
5. Research Objectives and Methodology 16
    5.1. Extracting a Time Series from an Audio Recording 16
    5.2. Finding the Length of the Curve 17
    5.3. Determining the Fractal Dimension 17
    5.4. Exploring Other Fractal Characteristics 17
6. Results 19
7. Analysis 22
    7.1. Lag Selection 22
    7.2. Fractal Dimension Distribution by Genre 23
    7.3. Predicting Genre by Fractal Analysis 24
8. Conclusions and Future Work 26
    8.1. Conclusions 26
    8.2. Future Work 26

References 28

1. Introduction to Fractals

In most engineering applications of mathematics, systems are created for sets with some type of regularity. This means that many sets are not included, nor can they be classified and studied in similar ways to the regular ones. This may mean studying differentiable functions, or systems with error that can be modeled with a normal distribution. In contrast with systems and tools considering variation to be the result of known variables and random variables, some systems exhibit chaotic properties and are best studied with tools that consider self-symmetry. Objects that are completely chaotic in this way are called fractals, and while a true fractal only a mathematical object, using fractals to estimate other systems can often be much more effective than traditional tool. For sets that resemble fractals, a value can be extracted from the set called the fractal dimension. This dimension serves as a fingerprint of the system, revealing the nature of the self-symmetry present [7]. In 2012, Tong and Chimka used fractal analysis of time-series temperature data to develop a fractal-based statistical quality control methodology [13]. This brings the notion of a system fingerprint into a useful light, as changes in the fractal dimension may indicate changes or failures in the system.

There are varying ways to define fractals. Here, we will use the list of common fractal properties provided by Falconer [6].

A **fractal** is some object F with the following properties

(1) F has a fine structure, i.e. detail on arbitrarily small scales.
(2) F is too irregular to be described in traditional geometric language, both locally and globally.
(3) Often F has some form of self-symmetry, perhaps approximate or statistical.
(4) Usually, the 'fractal dimension' of F (defined in some way) is greater than the topological dimension.
(5) In most cases of interest F is defined in a simple say, perhaps recursively.

While this list provides a basic framework for understanding fractals, it does not provide a precise mathematical definition. Part of the definition relies on the 'fractal dimension' which is not yet clearly defined. Many mathematicians have held more specific definitions, which often rule out sets other mathematicians consider fractal. This definition is much closer to a biologists definition of life. We have a list of properties that applies to most living things, but there are certain living objects that do not conform to every requirement [6]. It is easiest to understand these characteristics in geometric figures, where fractals were first recognized. Here are a few examples:

1.1. **The Cantor Set.** The Cantor set F is constructed by taking the interval $[0,1] \in \mathbb{R}$ and performing subsequent deletions. This deletion is the open middle third of all segments in the set. Let E_0 be the starting segment, $[0,1]$. We define E_k as E_{k-1} with the inner third of all segments removed. This means E_k consists of 2^k segments, each of length 3^{-k}.

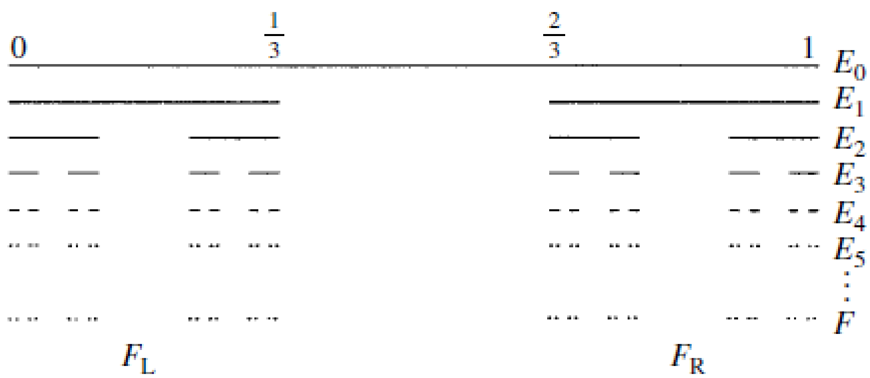

FIGURE 1. A graphic representation of creating the Cantor set. At each level, the middle third of each segment is removed. Notice that if the set is divided into two, each half is similar to the whole, although a third of the size [6].

Altogether, we have that $F = \cap_{k=0}^{\infty} E_k$ or $F = \lim\limits_{k \to \infty} E_k$. A visual representation of the Cantor set is shown in Figure 1. The Cantor set has some interesting properties. For instance, while there are uncountably many real numbers in F, it has a length of 0 using most conventional measures. Consider the five properties used to define a fractal where the set F is now the Cantor set.

- F has a fine structure, i.e. detail on arbitrarily small scales.

  For the Cantor set, it is impossible to "zoom in" enough to observe the finest level of detail. There is not a final level of E_k.
- F is too irregular to be described in traditional geometric language, both locally and globally.

  As mentioned prior, the segments of the Cantor set have no measurable length. This is true for the length of individual segments as well as their sum.
- Often F has some form of self-symmetry, perhaps approximate or statistical.

  Each level of the Cantor set has the same structure. Consider a segment on any level, say $[0, \frac{1}{9}]$ from E_2. The subsequent deletions behave identically to the deletions from the original segment $[0,1]$, forming geometrically identical set.
- The 'fractal dimension' of F (defined in some way) is greater than the topological dimension.

  This will be considered in following sections.

- The set F is defined in a simple, recursive way.

  This is apparent as removing the inner third is a recursive and straight-forward process.

1.2. **The Koch Curve.** The second example of a common fractal is the Koch curve, now becoming F. This figure is remarkably similar to the Cantor set, but instead of deleting the middle third of each segment, the middle third is replaced with two segments that would form an equilateral triangle with the removed segment. In a similar way to defining the Cantor set, we will define each level of the fractal's creation as E_k and the final set $F = \lim_{k\to\infty} E_k$. In Figure 2 there is graphic representation of this process.

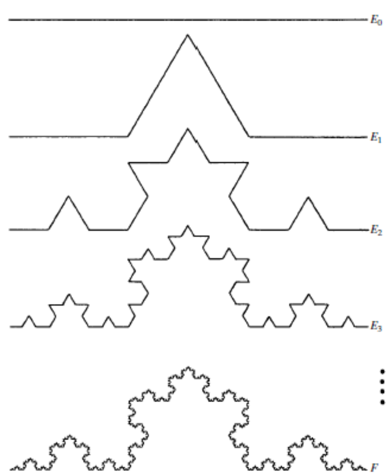

FIGURE 2. A graphic representation of creating the creating the Koch curve. With each level, the middle third is replaced with two lines that would form an equilateral triangle point upward from the missing segment [6].

As can be seen, the Koch curve resembles a snowflake. As k becomes larger, each level begins to affect only finer and finer detail. The polygonal curve E_k begins to approach the limiting curve, F [6]. This limiting curve cannot be analyzed by conventional means, however, as it is too irregular to construct traditional tangent lines. The length of any of the polygonal curves, E_k can be found to be $(\frac{4}{3})^k$. This clearly shows that F has infinite length. Let us again consider how this set is related to the definition of a fractal.

- F has a fine structure, i.e. detail on arbitrarily small scales.

  With each new level, the curve gains finer and finer detail and complexity. This continues to arbitrarily small scales.

- F is too irregular to be described in traditional geometric language, both locally and globally.

  While F is a curve, it is impossible to find tangent lines. Furthermore, the curve has endpoints and yet has infinite length.

- The set F is self-symmetric.

  The segment at any level of the curve produces it's own Koch curve as the levels progress to infinity.

- The set F is defined in a simple, recursive way.

  The definition is geometrically straight forward and is the product of recursive iterations.

The Koch curve holds these properties is a very similar way to the Cantor set. They are both perfectly self-symmetric, meaning that each small section of the set is geometrically symmetric to the full set. In the definition of a fractal, only statistical or approximate self-symmetry is required. In the traditional Koch curve construction, the added segments are always added in the same direction (either up from the original line or down). Consider the curve formed when this direction is determined by an independent random event at each level, say the tossing of a coin. In Figure 3 there is an image of a curve formed in this manner.

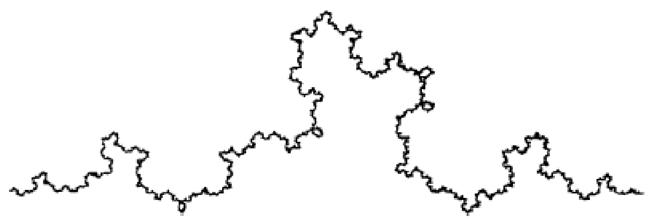

FIGURE 3. An example of a Koch curve where instead of always replacing the missing segment with two on the top. Approximately half of the time it is replaced on the bottom, making a fractal which is not completely self-symmetric [6].

This curve, like the first two fractals, has arbitrarily small complexity, cannot be studied by conventional methods, and is simply and recursively defined. The only difference is that it is not completely self-symmetric. This is why the definition is expanded to contain objects with statistical self-symmetry, as object like this curve clearly belong in the set of fractals.

1.3. **The Sierpinski Triangle.** The Sierpinski Triangle starts with an equilateral triangle. Similar to the Cantor set, subsequent deletions will be taken at each level to arrive at the final object. At each level, the triangle formed by the midpoints of each edge is removed for each triangle present as seen in Figure 4. This creates three triangles, each is one fourth the area of the starting triangle. Let E_i be the i-th iteration of this process. Similar to previous examples we have that $F = \lim_{i \to \infty} E_i$. In this case E_i is composed of 3^i triangles of area $\frac{1}{2} \cdot 4^{-i}$ assuming the original side length is 1.

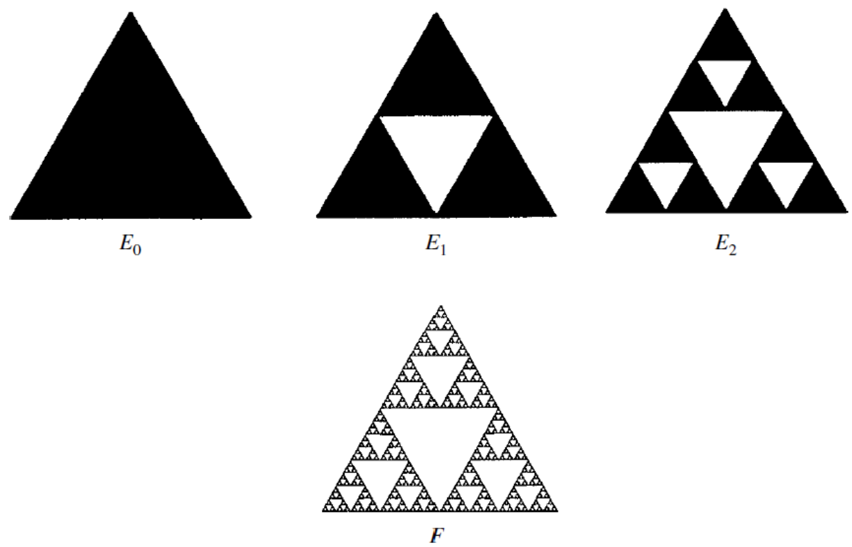

FIGURE 4. This graphic representation of creating the Sierpinski triangle shows the steps to create the figure. Notice each triangle of E_1 is a Sierpinski triangle itself, scaled to one half the original size[6].

The area of the starting triangle is $\frac{1}{2}$. For each iteration, the area is $\frac{3}{4}$ times the area of the previous figure. This means the area is equal to $\lim_{i\to\infty} \frac{1}{2} * (\frac{3}{4})^i$ or 0.

At this point it should be clear that the Sierpinski triangle is irregular in the same ways as the previous two examples.

2. The Fractal Dimension

Earlier it was mentioned that standard systems cannot adequately study fractal objects. Let us consider the examples. In the Cantor set, we have a set composed of many segments, however the length of the set is zero. Similarly, the length of the von Koch curve is infinite, although it has two endpoints and no area. The Sierpinski triangle also has zero area, but is constructed from an equilateral triangle. The traditional measures for objects in a specific dimension fall short to describe fractal objects in any meaningful way.

In a way, these objects seem to be between dimensions, but what could that even mean? Dimension does not indicate the amount of space something takes up, but the type of space it occupies. The most useful way for conceptualized the type of space between the dimensions is to consider the scaling properties of dimension. Consider two line segments, one being twice the length of the other. If squares are constructed with sides composed of each of the two. The longer segment produces a square that is four times the size of the smaller square. Similarly, if cubes are constructed from them, the larger cube is eight times the size as the smaller one. The dimension of each object perfectly describes this scaling relationship. The size of the final object, m, equals the scale, λ of the original object raised to the power of the dimension, D. We then have the following equation:

$$m = \lambda^D$$

or

$$D = \frac{\log m}{\log \lambda}$$

This concept can be extended to geometric fractals that are perfectly self-similar. Consider the Cantor set. A full Cantor set starting from any segment is equivalent to half of a Cantor set where the starting segment is three times the length. This would result in a dimension of log(3)/log(2). The von Koch curve formed from any starting segment is one fourth the size of a von Koch curve formed by a segment three times its length. This results in a dimension of log(4)/log(3). In general, we see in Figure 5 how these examples compare to each other and the whole-number dimensions.

2.1. **The Hausdorff Dimension.**

The Hausdorff dimension is defined for any set and is the most elegant fractal dimension that exists. It essentially captures the relationship between the maximum size of subsets used to cover the set and the number of sets required. Consider trying to cover a line segment with circles. If the circles are limited in size by one half their original size, it will take twice as many of

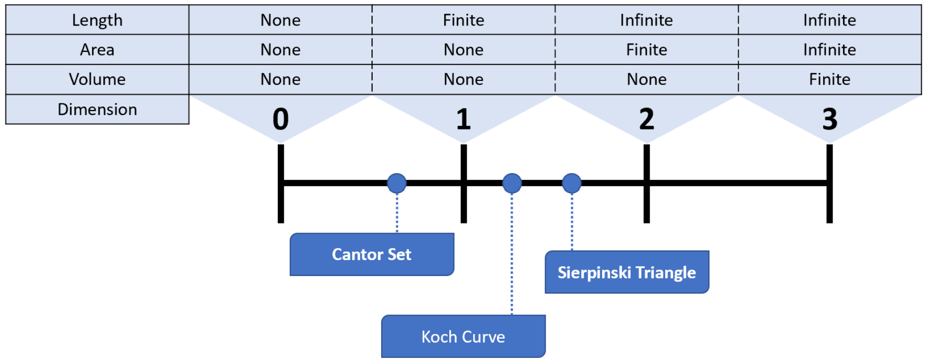

FIGURE 5. A diagram of the dimesional properties of some Euclidean dimensions. Notice that the fractal objects classified do not conform to all of the properties of one whole-numbered dimension.

them to cover the line segment. However, if the object is a square instead of a line segment, it will take four-times as many half-sized circles to cover the object. This new direction of thinking about dimension is applied in a very rigorous fashion to calculate the Hausdorff dimension.

The first step to building the Hausdorff dimension is the Hausdorff measure, which has some mathematical background. In Euclidean spaces, a measure is defined to be a function from some $\mathbb{R}^n$ to $[0, \infty]$ with some special conservation properties. Most readers will know the Lebesgue measures for at least three dimensions of space, length, area, and volume. In general, we refer to the n-dimensional Lebesgue measure as n-volume. The Hausdorff measure is an extension of Lebesgue measures by not limiting the measurement to a single, whole-numbered dimension. Instead, dimension has its own variable. If the dimension variable is set to 1, then the Hausdorff measure will be the same as length. If it is set to 2, then the Hausdorff measure will be the same as area and so on. In contrast with Lebesgue measures, Hausdorff measures are defined for "dimensions" that are not whole numbers.

The Hausdorff measure is defined by using a cover of the original set. This is a collection of sets whose union contains the original set. We also limit the size of each set in the cover with δ, the greatest distance allowed between elements in the set (or diameter of the set). Here we see a more formal mathematical definition:

Definition 2.1. Let F be a subset in $\mathbb{R}^n$ and $s \geq 0$. For any $\delta > 0$ we define

$$\mathcal{H}^s_\delta(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : U_i \text{ is a } \delta\text{-cover of } F \right\}.$$

We have that $\mathcal{H}_\delta^s(F)$ is the smallest possible sum for the s power of diameters of δ-covers of F. Also, the set of possible δ_1-covers is contained in the set of possible δ_2-covers when $\delta_1 < \delta_2$. Therefore, for $\delta_1 < \delta_2$, $\mathcal{H}_{\delta_1}^s(F) \geq \mathcal{H}_{\delta_2}^s(F)$. Thus as δ gets closer to zero, the $\mathcal{H}_\delta^s(F)$ approaches a limit. The **s-dimensional Hausdorff measure** of F is defined as follows:

$$\mathcal{H}^s(F) = \lim_{\delta \to 0} \mathcal{H}_\delta^s(F).$$

For readers without a strong mathematical background, it is only important to understand the following properties of the Hausdorff measure:

(1) When s is a positive integer, $\mathcal{H}^s$ is equivalent to s-volume as a Lebesgue measure
(2) For $s_2 > s_1$, $\mathcal{H}^{s_2} \leq \mathcal{H}^{s_1}$
(3) There is at most one value of s where $0 < \mathcal{H}^s < \infty$

As a result of these properties, we can build the Hausdorff dimension. We know that for some s in the positive real numbers that for all $t < s$ we have that $\mathcal{H}^t = \infty$ and for all $t > s$ we have that $\mathcal{H}^t = 0$. Notice while $\mathcal{H}^s$ can be infinity, zero, or some finite value between them, it is the location where this changes. This value of s, where $\mathcal{H}^s$ switches from infinity to zero is the Hausdorff dimension. Think of the Hausdorff dimension as a test for each s of whether the set occupies more or less space than an s-dimensional object.

2.2. **Box-Counting and Other Fractal Dimensions.** While the Hausdorff dimension is a very elegantly captures the goals of finding the fractal dimension. It is not very easy to apply to fractal objects, particularly ones that are not perfectly self-similar. For that reason, there are other methods that attempt to approximate the efforts of the Hausdorff dimension to estimate the fractal dimension of fractals and approximate fractals. The most notable of these is the box-counting dimension, a method applied to objects that can lie in a two-dimensional space. In this method, the fractal set is superimposed by a square grid. Then the number of grid boxes the figure overlaps with is recored along with the side length of the grid size. Then, the boxes are reduced in size and the process is iterated. Looking back at the similarity dimension, where $m = \lambda^D$, we can update m and λ to the new context to solve for D. While before λ was the scaling factor, now λ is the side length of the boxes. Similarly, while m was the final size of the object, m_λ is the number of λ-sized boxes required to cover the object. In its most appropriate form, the box-counting dimension of a figure follows:

$$D = \lim_{\lambda \to 0} \frac{\log m_\lambda}{-\log \lambda}$$

There is a negative sign introduced to equation because of the interpretations of m and λ. For practical purposes, the box-counting dimension is often estimated as the slope between $\log m_\lambda$ vs. $\log \lambda$ for observations from many different sized λ. These calculations extend the previous formula into the following relationship:

$$D \sim \frac{\log m_\lambda}{-\log \lambda}$$

In Figure 6 we see the box-counting process being applied to the Koch curve. Using the data obtained from these three iterations, we could begin to estimate the fractal dimension of the curve. Notice that this process would still provide a fractal dimension, even for objects that are not true fractals until the granularity of the grid was smaller than the complexity of the figure.

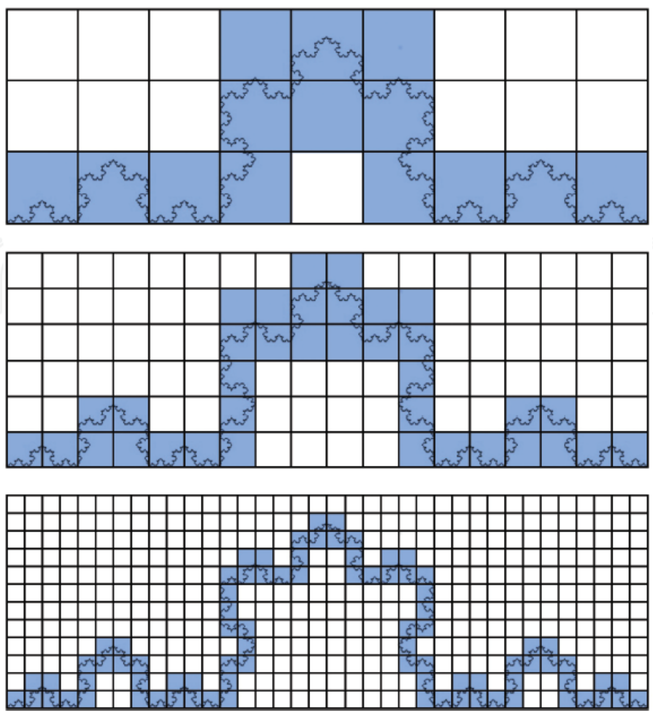

FIGURE 6. Applying the box-counting method to the Koch curve. Notice that as the grid gets more fine, the number of boxes required to cover the curve increases while the area required to cover the curve decreases. The number of boxes required is inversely correlated with the side length of the boxes raised to the box-counting dimension of the curve. [11].

There are many other ways to estimate the fractal dimension of objects. Many methods have been developed with slightly different goals and are often well-suited to a particular kind of set. While some are variations on the box-counting dimension, others are very different. Most reflect a variation of the dimension equation from Section 2.

3. Measuring the Fractal Dimension of Time-Series Data

There has been significant research into measuring the fractal dimension for time series data. These are sets where variation in the state of the system has self-symmetric properties. For example, stock price at a particular time is system-state variable. The stock price over time is in fact a time-series fractal object (to some extent). Remember that at some level of granularity, the fractal qualities will disappear, as stock prices cannot change over arbitrarily small periods of time at arbitrarily small levels. There are other types of time-series data that have been found to have fractal qualities including pressure in oil pumping and music.

While properties such as length and area are concerns for geometric figures, time-series data is usually more concerned with mean and variance of the system-state variable. For a true fractal time-series, these may not exist, making the fractal dimension a very important identifier for the series [9]. In contrast with geometric fractals, time-series fractals are based on measurement with one scale and time measured at some scale. Unlike a curve in $\mathbb{R}^2$, these scales are independent of each other. This means the time-series curve has no specific shape. As such, conventional methods are not always effective for time series data. To consider the data the same way as a curve (such as the Koch curve) has been found to produce inconsistent results. Changes in the scaling produces varying results when applying methods such as box counting.

In a paper by Suleymanov, the length of the curve is not considered with respect to time [1]. The time series is considered a set of measurements over an interval of time. The overall length of the curve is considered to be the sum of the absolute deviations between each subsequent measurement. The fractal dimension is found by increasing the lag between measurements used. The overall length of the curve is compared with the number of measurements used to find it. In this method we consider both the number of observations used to determine curve length k, operating as our λ, and the overall length of the curve L, operating as our m. With adjustments made for the definitions of λ and m, we have the following:

$$D - 1 \sim \frac{\log k}{\log L}$$

There have also been several other methods implemented for determining the fractal dimension of a time-series. The Suleymanov method was found to be very reliable when compared with other methods and provided a large computational advantage [1]. The algorithms are also easily implementable into MATLAB, which was used for this research. To avoid confusion, the fractal dimension found using the

Suleymanov method will be referred to as the Suleymanov dimension. Here is a more explicit framing of the method. In order to estimate D, we must find several observations of L and k, that is the length of the curve when a different number of observations are used to determine the length. Let the time series A be the sequence of values y_i where $i \in [1, 2, ..., n]$. These are assumed to be taken at equally distant periods of time. We define j as the lag interval used for the measurement, which is tied to a unique value for k and L. With each j, we first find k, the number of intervals used to measure the length. If every observation is used (meaning $j = 1$), then $k = n - 1$. In general, if every j observations are used we have that $k_j = \lfloor \frac{n-1}{j} \rfloor$.

We define L_j, the length of the curve using the lag interval j as follows:

$$L_j = \sum_{i=1}^{k} |y_{(i*j)} - y_{(i*(j-1))}|$$

Altogether, we see that L is the sum of absolute differences. The elements of A used are only ones that are multiples of j. Thus L is the total length using a lag interval of j. When L_j is plotted versus k_j at various step lengths, we can find the slope of the line of best fit. Based on our earlier definition of the Suleymanov dimension, this is equivalent to $D - 1$, thus providing D.

4. Measuring the Fractal Dimension of Music

Many researchers have worked on studying the fractal dimension of music in various forms. Because the fractal dimension is an identifying mark of a fractal set, understanding the fractal dimension of music may unlock a deeper understanding of musical characteristics. One potential avenue is the classification of music. In studies of traditional Indian music, the fractal dimensions from music in three different genres were found to be statistically different from each other. There is also a large amount of research into the creation of music with machine learning algorithms. An understanding of the fractal qualities of different types of music may offer a new method for music creation.

The type of object used for fractal analysis of music has varied significantly across the research in the field. Some research has been oriented at the written structure of musical compositions. In 1994 Stephanie Mason used fractal systems to generate melodies that conform to ”Western melodic expectations” by utilizing the self-symmetry of musical structures [10]. In contrast, in 1987 Klimontovich investigated the fractal properties of the ”long-term dynamics” of a piece from an audio recording[8]. In this study, the time series of what is called the *audio variable* was used for the analysis. It was similarly found that this object has fractal self-similarity for musical recordings and that it could be a useful tool for objective characterization of music and a potential basis for music composition.

This audio variable used could be described in many ways, but ultimately is the information needed to recreate the sound of the recording. This is a time series of values captured by a digital recorder from a microphone, which can then be sent to a speaker to reproduce the sound. When normalized, this series takes values between -1 and 1. It is easiest to conceptualize with the example of a speaker, in which each value corresponds to a position of the diaphragm of the speaker. Think of 1 as being fully pushed out and -1 as being fully pulled in. In order to distort the air in the precise way to recreate the recording, the speaker diaphragm must move between these extremes in accordance with the time series. It is common for recordings to have a sampling rate of 48 kHz, meaning that there are 48,000 values in the time series per second of the recording. The higher this rate, the higher the fidelity of the recording, as it will more closely resemble the audio it is capturing. The time series of this audio variable is the type of musical object that this research is primarily considering.

In 2000, Bigerelle revisited and tested whether this audio variable could be used to distinguish between musical recordings [2]. It was confirmed that fractal dimension could be used to discriminate music based on its dynamic aspects. Since then, researchers Das and Das have published many papers working to classify music based on fractal dimension, usually working with traditional Indian songs. First they used fractal analysis to distinguish between 3 genres of Indian songs [3]. Next, they began to evaluate the effects of instrumentation [4]. They have even evaluated whether the skill level of vocal performers affects the fractal dimension of the recording [5]. In each of the cases, any accompaniment part of the recording was eliminated, leaving only a solo performance. Additionally, the data used was converted to ASCII format, possibly leading to a different analysis of the time series. Altogether, they found fractal dimension to be significant between the groups studied.

Altogether, the fractal dimension has been found to be a very interesting characteristic of music. Research has begun to examine the methods for categorizing music with this measure and has begun to find its utility. Overall, there is more research to be done with the types of music considered, the method employed for finding the fractal dimension, and the ways the fractal dimension is used to characterize the music.

5. Research Objectives and Methodology

Much of the current research has only applied fractal analysis to short clips of music due to methods that are computationally intense. Using a method of computation that is more efficient opens the door to a larger scale investigation of the fractal qualities of music. The goal of this research is to apply methods of determining the fractal dimension of time series data to a selection of musical recordings. This is then used to measure if there is statistical difference in the fractal qualities of musical recordings from different genres. Ultimately the results are used to build a simple machine learning model to evaluate the utility of the Suleymanov dimension to classify music by genre.

The audio recordings chosen for this research come from a dataset that is prominent in musical genre classification from a paper by Tzanetakis and Cook published in 2002 [14]. This dataset contains 1000 audio tracks, each of 30 seconds in length. There are 10 genres represented, each with 100 tracks recorded at 22,050Hz. The dataset was originally used to test a method of automated music categorization by genre. It serves a good comparison for how well the Suleymanov dimension can categorize music.

5.1. Extracting a Time Series from an Audio Recording.

Previously, it was mentioned that this research would construct the time-series data from a variable called an audio variable. Adding a litte more detail, this variable most directly represents the micro changes in air pressure due to the traveling sound waves. This change in pressure can be captured with a microphone and digitally stored as a time series. When normalized to the range $[-1, 1]$, it can then be reproduced when a speaker imitates the original sound with vibrations that create new sound waves in the air.

This time series carries all the information of the sound waves. Consider a time series where the values resemble a sine wave such that there are 440 cycles a second. If this time series were to be played through a speaker, we would recognize the result as a single pitch, in this case ”A” and the amplitude of the wave would be the volume. The nature of musical recordings are much more complex than simple sin waves, however, due to the presence of overtones, harmonies, and timbre (the word capturing the different qualities between instruments and tone). It is because of this complexity and self-symmetric properties that music is conducive to fractal analysis.

From any mp3 file of a musical recording, free online software can be easily utilized to generate a wav file of the recording. This file format can be read by the MATLAB function *audioread* to store the audio variable as a time series. At this point, a few operations were performed to prepare the files for analysis, which are listed here:

(1) Each of the channels were combined into one time series. If the recording was stereo, then the left and right channel were averaged together, thus creating a mono recording
(2) All the leading and trailing ”zeros” in the data were removed so that the time before and after the music occured would not interfere with the results

While each recording may have been normalized to a different level, it was found that the fractal dimension of a recording was not affected by its scaling, so normalization of the recordings was not necessary.

5.2. **Finding the Length of the Curve.** The length of the curve is highly dependent on the number of observations used to calculate the length. It is important to remember that the true sound variable is continuous, given any time it has a value. Because of the complexity of music, the variation of the variable is not fully captured by a time series that only collects information at discrete locations. The length of the curve was found using the methodology introduced by Suleymanov. In this method, L_j is the sum of the absolute differences between every j-th observation of the time series. We always have that L_1 is the greatest possible length with the given granularity of data. In this example, we also know that L_2 would be the greatest length if the sampling rate was only half (meaning observations only occurred half as frequently). Given the the extracted variable from the previous step, a MATLAB function was created to find the length of a curve given a time series and a step size. Also, the total number of smaller lengths used was output, which is k_j and is used for calculating the fractal dimension.

5.3. **Determining the Fractal Dimension.** The fractal dimension used will be the Suleymanov dimension, as discussed above, and lies in the relationship between the length of the curve and the number of observations used to measure the length of the curve. Because the observations do not represent a true fractal and there is a limit to the granularity contained in the time series, we can estimate D with the slope between $\log(L)$ and $log(k)$ within a range where the time series resembles a fractal. A MATLAB function was created to output the Suleymanov dimension given a series for L and k.

5.4. **Exploring Other Fractal Characteristics.** Because of the time series has a limited granularity, the fractal dimension is limited to a range where complexity is evident. If the Koch curve was generated recursively for any finite number of steps, the fractal relationship would decay as the measurements became more precise than the complexity of the curve. In addition to outputting the fractal dimension, a MATLAB function was created to determine if there is a step size where the complexity of the curve is lost, having only a Suleymanov dimension of 1, which corresponds to a slope of zero.

In this function, the set of L and k are divided into two at a splitting point. On the left the slope is determined with linear regression and on the right the slope is assumed to be zero and are thus estimated with their mean. The function moves this splitting point between each of the points and selects the splitting point minimizing the overall total error.

The outputs of this function are the number of observations belonging to the part with zero slope and the Suleymanov dimension using only the other observations. Altogether, the MATLAB functions output the Suleymanov dimension using the lag intervals chosen, the number of lags where the slope is zero, and the Suleymanov dimension using the observations not included in from the zero-slope component. Each of these three are shown in Figure 7.

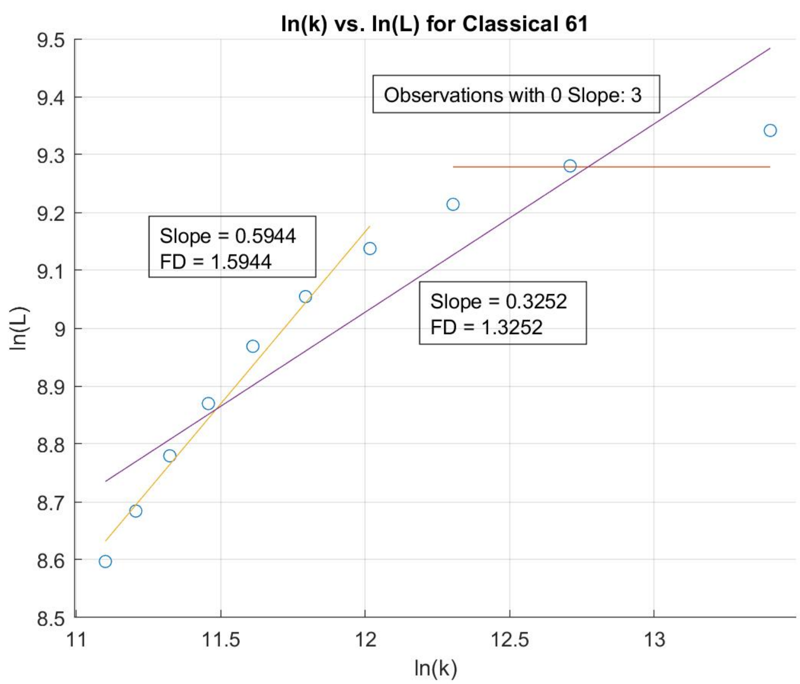

FIGURE 7. This is a plot of $\ln L$ vs. $\ln k$ for the classical recording number 61. The Suleymanov dimension calculated across all lag intervals is 1.325. However, the 3 smallest lag intervals more closely resemble having no slope. When these are removed, the Suleymanov dimension calculated is 1.594.

6. RESULTS

The fractal analysis described above was applied to the 1,000 thirty-second audio files. There are ten genres represented: blues, classical, country, disco, hip hop, jazz, metal, pop, reggae, and rock. Each of these genres has 100 recordings in the dataset. The titles of these recordings were collected by Dr. Sturm and can be found at this link http://www.eecs.qmul.ac.uk/ sturm/research/GTZANindex.txt [12]. Twenty lag intervals were used for plotting $\ln(k)$ vs. $\ln(L)$ such that $j \in 1, 2, ..., 20$. The following table shows some example data produced by the analysis, where only song number 1 was selected from each genre. The table shows the genre, the song identifier, the Suleymanov dimension, the r^2 value for estimating $\ln(k)$ vs. $\ln(L)$, the adjusted r^2 value, the number of lags found with zero slope, the Suleymanov dimension taken without these observations, and the squared error of piece-wise estimation for $\ln(k)$ vs. $\ln(L)$.

Genre	**Song**	**FD**	r^2	r^2_{adj}	**cutoff**	**FD$_2$**	**error**
Blues	1	1.40	0.98	0.98	1	1.43	0.02
Classical	1	1.58	0.79	0.77	6	2.25	0.10
Country	1	1.74	0.98	0.98	1	1.81	0.02
Disco	1	1.82	0.97	0.97	2	1.95	0.09
Hip Hop	1	1.78	0.99	0.99	1	1.83	0.03
Jazz	1	1.43	0.94	0.93	3	1.58	0.03
Metal	1	1.78	0.99	0.99	1	1.84	0.03
Pop	1	1.67	0.99	0.99	1	1.64	0.01
Reggae	1	1.43	0.97	0.97	2	1.51	0.03
Rock	1	1.41	0.97	0.97	2	1.49	0.01

Using this data for all 1,000 audio files, we can understand the fractal properties of the recordings based on the genre. In Figure 8 we see the distributions of the Suleymanov dimension across genre.

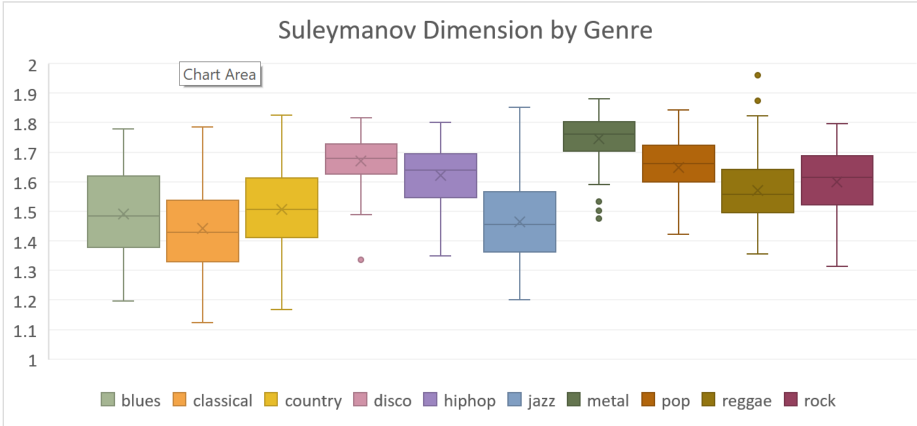

FIGURE 8. Here we see the box plot of the Suleymanov dimensions for each recording analyzed separated by genre.

While each genre seems to contain a different distribution, they do not seem to all have separate fractal dimensions. From here, one-way ANOVA testing was performed to test whether genre was significant with regard to the Suleymanov dimension. The results of this test are shown in the next table.

	SS	**DF**	**MS**	**F**	**p**
Model	8.80	9	0.98	69.29	0.0001
Error	13.98	990	0.01		
Total	22.78	999			

Because the p-value is very low, we see evidence that FD is not independent of genre. A Tukey grouping was then applied to determine the significance of the differences between genres. The least significant difference of FD was found to be 0.0533 and thus there were 6 groupings of genres with no statistical difference in Suleymanov dimension. These can be further broken down into 3 divisible groups. These results are seen in Figure 9. Metal was the only genre in which the mean Suleymanov dimension was statistically different from all other genres. The testing was also applied to the adjusted Suleymanov dimension, which is seen on the right of the figure.

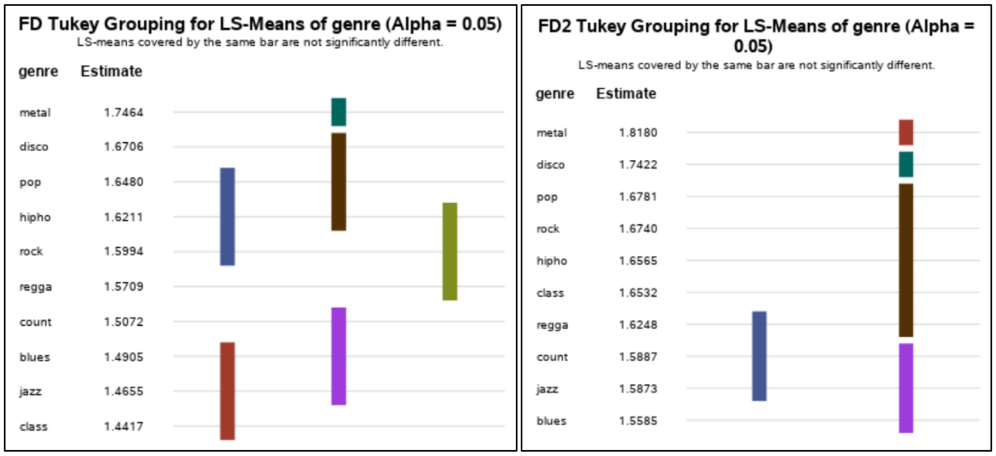

FIGURE 9. On the left side, we see the mean FD for each genre. Genres connected by a bar indicate there is not statistical difference between them. On the right, we see the adjusted FD where observations with zero slope are removed.

Additionally, the six values recorded in the raw data table were used in MATLAB's "Classification Learner" add-in to predict genre with machine learning methodology. An array of techniques were employed with 5-fold validation to determine how valuable these six measures are for predicting the genre of an audio recording. The best results came from a k-nearest neighbor algorithm, reaching an accuracy of 35.8%. In Figure 10 there is a confusion matrix showing the true-positive rates across genres.

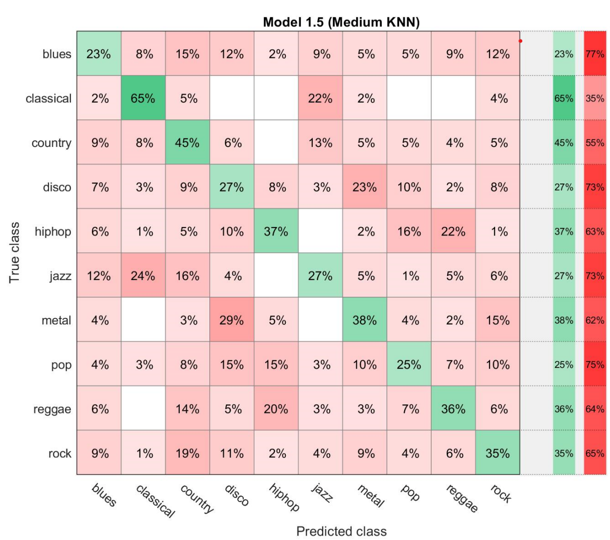

FIGURE 10. This is the confusion Matrix from the medium K-Nearest Neighbor (KNN) model. Each value represents the percent of songs predicted to belong to a specific genre given the real genre, so they sum to 100% across the rows. The best performer was classical. Out of the 100 classical songs, 65 were correctly predicted based on the fractal qualities.

7. Analysis

7.1. **Lag Selection.** One of the most important decisions while using the Suleymanov dimension is the selection of lag intervals, or j when the Suleymanov dimension is described in Section 3. This value is an integer as small as 1, and can increase such that it remains small compared to the length of the time series. It was discovered that the selection of j is very relevant to the calculation of the Suleymanov dimension. In Figure 11 we see this with the familiar example, classical piece number 61. Here, $\ln(L)$ is plotted versus $\ln(k)$ with lag intervals at different orders of magnitudes. The ten right-most points increase in lag intervals by 1. Then the lag intervals from 10 to 100 are represented at steps of 10. Finally, the lag intervals from 100 to 1,000 are represented by steps of 100. There are 28 total lag intervals represented.

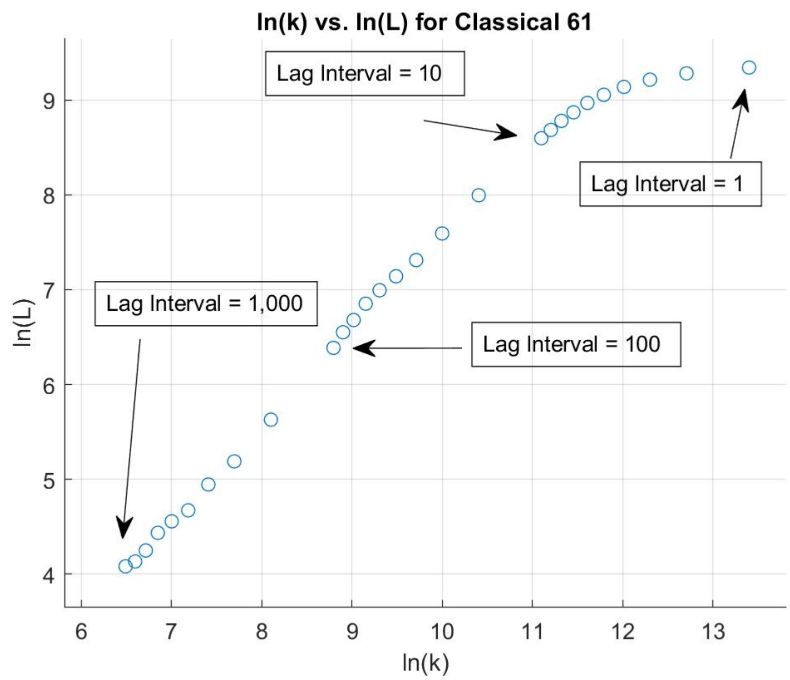

FIGURE 11. The plot of $\ln(L)$ vs. $\ln(k)$ for many lag intervals. The slope between lag interval 10 and 100 is very similar to the slope between lag intervals 100 and 1,000. This slope begins to approach zero as the lag interval gets closer to 1.

Notice that after the first few observations (on the top right corner) the slope becomes very stable. This shows consistency in the fractal dimension of the time-series until the granularity becomes too fine. This means that the complexity of the time series diminishes at a certain precision. This raises questions on the appropriate approach to measure the fractal dimension. Should these observations be considered or ignored? Furthermore, what type of characteristics could be measured by this decrease in slope? In response to these questions, the data was collected in a manner that attempted to balance the affects of including or excluding the data. First, the lag intervals were chosen to increase by 1's from 1 to 20. For almost all files, the slope of the regression began to converge well for lags greater than 10. Furthermore, using methodology discussed earlier, the lowest lags were considered for removal for a second fractal dimension calculation. It is important to remember that the number of large lags included affects the number of cutoff lags. While this research chose to include twenty total lag intervals, the results are affected by this decision. Choosing more lags or lags of greater value would likely affect the cutoff selected for excluding observations. The number of excluded lags was recorded as well as the modified Suleymanov dimension. These measures work together to provide a more rounded view of a recording's fractal qualities.

7.2. **Fractal Dimension Distribution by Genre.** In the results section, we saw that average fractal dimension within a genre ranges from 1.44 for classical recordings to 1.75 for metal recordings. In the adjusted Suleymanov dimension, FD2, we see in increase in the fractal dimension of anywhere between 0.03 to 0.21 with an average difference of 0.08. This tells us that not only is the decision of lag intervals important, but that the affects of the lag decisions vary across genres. The genre with the most change was classical, where on average recordings lag intervals up to 3.64 excluded due to their decreased slope. This placed classical music as the sixth largest adjusted fractal dimension, even though it had the smallest fractal dimension before the exclusions.

The results show that after exclusions classical recordings have a statistically larger Suleymanov dimension than country, jazz, or blues recordings. This means the curve of $\ln(L)$ vs. $\ln(k)$ for classical music begins steeper, but has a sharper drop-off as the lag intervals become small. This means the time series has more self-symmetric complexity, but it is not continued at the fine level.

The genre with the largest average Suleymanov dimension, both before and after the adjustment, was found to be metal. This is expected to be because of the nature of the audio content. The metal genre is characterized by having a very full sound space and heavy distortion. While other genres may have more space between sounds and have limited frequencies present, metal is more likely to have sustained sound with many frequencies present from the distortion. These create a more complex wave-form of the audio. For these reasons, it is not surprising that metal music would have a larger fractal dimension than other recordings.

7.3. **Predicting Genre by Fractal Analysis.** Because it was found that the distinguishing features between genre included not only one measurement of slope, but the more complex relationship between $\ln(L)$ vs. $\ln(k)$, each of the six features recorded in the results section were included for predicting genre. Overall, the performance of the classification were poor, only achieving 35.8% accuracy. This is much less than the classifications that can be obtained as a result of audio signal processing. The same dataset was classified by these means with 61% accuracy by Tzanetakis in 2002 [14].

There is not data available about the performance of the Tzanetakis model at the song-specific level in order to assess the affects of adding in the fractal predictors. Because the data used for the classifications vary significantly, it is very likely that the inclusion of the fractal information would benefit a model using other audio features. In addition, the audio transformations used by Tzanetakis likely produces time-series with fractal characteristics that could additionally aid classification even more. Considering the results of the Tukey grouping, the classification rate makes sense. On average, a genre belongs to a group of 3.4 genres where there is not statistical difference, as seen in the following table:

Genre	**Count of Genres with Statistically Similar FD**
Metal	1
Disco	3
Pop	4
Hip Hop	5
Rock	4
Reggae	3
Country	3
Blues	4
Jazz	4
Classical	3
Average	**3.4**

Using only the FD, it would be unlikely for a model to discriminate between genres where there is no statistical difference, leading to an expected performance of $\frac{1}{3.4}$ or about 28% accuracy. This assumes it can perfected discriminate between statistically different genres and not at all between those with no statistical difference. Because these are not completely true observations and there is additional information available to the models, the accuracy was able to be increased to almost 36%.

8. Conclusions and Future Work

8.1. **Conclusions.** The goal of this research was to extend the work on fractal analysis in the field of musical study. It applied a method by Suleymanov for finding the fractal dimension of a time series to the raw audio variable contained in a digital recording. Functions were successfully created in MATLAB to extract the Suleymanov dimension from audio files with the ability to alter the desired lag intervals.

The methodology was then adapted to further distinguish the fractal qualities of the recordings when the lag intervals are small. The functions created were then applied to a database of 1,000 audio files, which held 100 30-second clips from ten different genres. The data was recorded for statistical analysis and genre prediction from fractal qualities.

One-way analysis of variance testing was performed and further Tukey grouping revealed three disjoint groups across the ten genres for statistically different fractal dimensions.

The full data was then used in MATLAB's "Classification Learner" to investigate the ability of the fractal analysis to predict genre. While the accuracy was only 35.8%, this avenue seems to be a profitable addition to current work on musical classification.

8.2. **Future Work.** This work has helped developed several questions across its domains. For the analysis of time-series fractal objects, the Suleymanov dimension was found to approach zero as lag intervals decreased in size. Further work could investigate other ways to characterize this curve rather the steps taken in this paper, linear regression and a piece-wise regression where the one component has no slope. Perhaps other parameters better characterize the relationship between $\ln(L)$ and $\ln(k)$. These other parameters may contain more identifying fractal characteristics than the parameters used in this research.

In the domain of musical genre classification, fractal analysis should be combined with other types of audio analysis to further improve the classification methods. Additionally, the fractal qualities of genres provide additional information to incorporate in algorithmic approaches to music generation. It is recommended that using the Suleymanov dimension could be further investigated in conjunction with other tools.

Finally, it is recommended that the Suleymanov dimension could be applied to a transformed version of the original audio files. In most other approaches to musical classification, methodology includes the decomposition and filtering of the raw audio, creating more meaningful data. These transformations likely hold information that also has fractal qualities, likely with more distinct values across genres.

References

[1] Aydin J. Ismaylov Arif A. Suleymanov, Askar A. Abbasov. Fractal analysis of time series in oil and gas production. *Chaos, Solitons and Fractals*, 41, 2009.

[2] M. Bigerelle and A. Iost. Fractal dimension and classification of music. *Chaos, Solitons and Fractals*, 2000.

[3] Atin Das and Pritha Das. Classification of different indian songs based on fractal analysis. *Complex Systems*, 15, 2005.

[4] Atin Das and Pritha Das. Fractal analysis of different eastern and western musical instruments. *Fractals*, 14, 2006.

[5] Atin Das and Pritha Das. Fractal analysis of songs: Performer's preference. *Nonlinear Analysis: Real World Applications*, 2010.

[6] Kenneth Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley, 1990.

[7] Jan W. Kantelhardt. Fractal and multifractal time series. *Mathematics of Complexity and Dynamical Systems*, 2008.

[8] Yu. L. Klimontovich and J.P. Boon. Natural flicker noise ("1/f noise") in music. *Europhysics Letters*, 1987.

[9] Ming Li. Fractal time series—a tutorial review. *Mathematical Problems in Engineering*, 2010.

[10] Stephanie Mason and Michael Saffle. L-systems, melodies and musical structure. *Leonardo Music Journal*, 4, 1994.

[11] Ian Pilgrim and Richard Taylor. *Fractal Analysis of Time-Series Data Sets: Methods and Challenges*. 11 2018.

[12] Bob L. Sturm. An analysis of the gtzan music genre dataset. *Proceedings of the second international ACM workshop on Music information retrieval with user-centered and multimodal strategies*, 2012.

[13] Jingjing Tong and Justin R. Chimka. Statistical process control for the fractal dimension of a time series. *Int. J. Quality Engineering and Technology*, 3, 2012.

[14] George Tzanetakis and Perry Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 2002.